

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2006, 20:53:35 ; Search time 46 Seconds
(without alignments)
666.797 Million cell updates/sec

Title: US-09-872-505A-6
Perfect score: 1892
Sequence: 1 MITMLWHAMPELNTARLMA.....PLAQREDDDDWDEDDW 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCUS COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1892	100.0	371	2	US-09-050-739-92
2	1873	99.0	368	2	US-08-818-112-114
3	1873	99.0	368	2	US-08-818-111-109
4	1873	99.0	368	2	US-09-056-556-114
5	1873	99.0	368	2	US-09-072-596-109
6	1873	99.0	368	2	US-09-072-967-114
7	1873	99.0	368	2	US-10-193-002-109
8	1873	99.0	368	2	US-10-084-843-114
9	366	19.3	400	2	US-09-073-009-126
10	366	19.3	400	2	US-09-073-010-126
11	344.5	18.2	394	2	US-09-712-363-205
12	330.5	17.5	729	2	US-09-223-040-2
13	330.5	17.5	729	2	US-09-287-849-2
14	328.5	17.4	596	2	US-09-287-849-26
15	326	17.2	423	2	US-09-073-009-142
16	326	17.2	423	2	US-09-073-010-142
17	326	17.2	710	2	US-09-287-849-16
18	326	17.2	856	2	US-09-287-849-12
19	318	16.8	600	2	US-08-818-112-107
20	317	16.8	391	2	US-08-818-111-102
21	317	16.8	391	2	US-08-818-111-102
22	317	16.8	391	2	US-09-056-556-107
23	317	16.8	391	2	US-09-072-596-102
24	317	16.8	391	2	US-09-072-967-107
25	317	16.8	391	2	US-10-193-002-102
26	317	16.8	391	2	US-10-084-843-107
27	312	16.5	396	2	US-08-818-112-111

28	312	16.5	396	2	US-08-818-111-106	Sequence 106, App
29	312	16.5	396	2	US-09-056-556-111	Sequence 111, App
30	312	16.5	396	2	US-09-072-596-106	Sequence 106, App
31	312	16.5	396	2	US-09-072-967-111	Sequence 111, App
32	312	16.5	396	2	US-10-193-002-106	Sequence 106, App
33	312	16.5	396	2	US-10-084-843-111	Sequence 111, App
34	311.5	16.5	1271	1	US-08-095-734-2	Sequence 2, Appli
35	311.5	16.5	1271	1	US-08-444-623-2	Sequence 2, Appli
36	311.5	16.5	1271	2	US-08-471-869-2	Sequence 2, Appli
37	311.5	16.5	1271	2	US-09-342-563-2	Sequence 2, Appli
38	311.5	16.5	1271	4	PCT-US94-08267-2	Sequence 2, Appli
39	294.5	15.6	359	2	US-08-818-112-109	Sequence 109, App
40	294.5	15.6	359	2	US-08-818-111-104	Sequence 104, App
41	294.5	15.6	359	2	US-09-056-556-109	Sequence 109, App
42	294.5	15.6	359	2	US-09-072-596-104	Sequence 104, App
43	294.5	15.6	359	2	US-09-072-967-109	Sequence 109, App
44	294.5	15.6	359	2	US-10-193-002-104	Sequence 104, App
45	294.5	15.6	359	2	US-10-084-843-109	Sequence 109, App

ALIGNMENTS

RESULT 1
US-09-050-739-92
; Sequence 92, Application US/09050739
; Patent No. 6641814
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDRING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/050,739
; CURRENT FILING DATE: 1998-03-30
; EARLIER APPLICATION NUMBER: 0376/97
; EARLIER FILING DATE: 1997-04-02
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-92

Query Match	100.0%	Score 1892;	DB 2;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 1.4e-154;		
Matches 371;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MITMLWHAMPELNTARLMAGAGPAPMLAAAGWQTLAALDAQAVELTARLNSLGEAWT	60	
Db	1	MITMLWHAMPELNTARLMAGAGPAPMLAAAGWQTLAALDAQAVELTARLNSLGEAWT	60	
Qy	61	GGGSDKALAAATPMVWLQTAQTRAMQATAQAAYTQAAATTPSLPEIAANHITQA	120	
Db	61	GGGSDKALAAATPMVWLQTAQTRAMQATAQAAYTQAAATTPSLPEIAANHITQA	120	
Qy	121	VLTAATFFGNTIPIALTENDYFIRMNQAALAMEVYQAEVAVNTLFEKLEPMASILDPG	180	
Db	121	VLTAATFFGNTIPIALTENDYFIRMNQAALAMEVYQAEVAVNTLFEKLEPMASILDPG	180	
Qy	181	ASQSTTNPIFGMPSPGSSSTPVGQLPPAAATQTLQGLGEMSGPMQOLTOPLOQVTSLSFQVG	240	

Db 181 ASQSTTNPIFGMPSPGSSTPVQQLPPAATQTGLGEMSGPMQQLTQPLQQVTSLSFQVG 240
Qy 241 GTGGGNPADEEAAQMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPMS 300
Db 241 GTGGGNPADEEAAQMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPMS 300
Qy 301 QLIEKPVAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQERED 360
Db 301 QLIEKPVAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQERED 360
Qy 361 DEDDWEDEDDW 371
Db 361 DEDDWEDEDDW 371
RESULT 2
US-08-818-112-114
; Sequence 114, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-112-114

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MLWHAMPELNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 63
Db 1 MLWHAMPELNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 60
Qy 64 SDKALAAATPVVWMLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
Db 61 SDKALAAATPVVWMLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

Qy 124 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
Db 121 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
Qy 184 STTNPIFGMPSPGSSTPVQQLPPAATQTGLGEMSGPMQQLTQPLQQVTSLSFQVG 243
Db 181 STTNPIFGMPSPGSSTPVQQLPPAATQTGLGEMSGPMQQLTQPLQQVTSLSFQVG 240
Qy 244 GGNPADBEAAQMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPMSQLI 303
Db 241 GGNPADBEAAQMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPMSQLI 300
Qy 304 EKPVPAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQEREDDED 363
Db 301 EKPVPAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQEREDDED 360
Qy 364 DWDEDDW 371
Db 361 DWDEDDW 368

RESULT 3

US-08-818-111-109
; Sequence 109, Application US/08818111
; Patent No. 6338852
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,111
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-111-109

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MLWHAMPELNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 63
Db 1 MLWHAMPELNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 60

Db 1 MLWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSIGEAWTGGG 60
 QY 64 SDKALAAATPMVWLOQTASTOAKTRAMQATQAATAAAYTOAMATTPLSLPEIAANHITQAVLT 123
 Db 61 SDKALAAATPMVWLOQTASTOAKTRAMQATQAATAAAYTOAMATTPLSLPEIAANHITQAVLT 120
 QY 124 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 183
 Db 121 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 180
 QY 184 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
 Db 181 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
 QY 244 GGNPADEEAAQWGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303
 Db 241 GGNPADEEAAQWGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300
 QY 304 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGSTRPGLVAPAPLAQEREDDED 363
 Db 301 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGSTRPGLVAPAPLAQEREDDED 360
 QY 364 DWDEEDDW 371
 Db 361 DWDEEDDW 368

RESULT 4
 US-09-056-556-114
 ; Sequence 114, Application US/09056556
 ; Patent No. 6350456
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
 ; TREATMENT OF SEQUENCES: 241
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED AND BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 07-APR-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.457
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 114:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 368 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-056-556-114

Query Match 99.0%; Score 1873; DB 2; Length 368;
 Best Local Similarity 99.7%; Pred. No. 6e-153;
 Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 MLWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSIGEAWTGGG 63

Db 1 MLWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSIGEAWTGGG 60
 QY 64 SDKALAAATPMVWLOQTASTOAKTRAMQATQAATAAAYTOAMATTPLSLPEIAANHITQAVLT 123
 Db 61 SDKALAAATPMVWLOQTASTOAKTRAMQATQAATAAAYTOAMATTPLSLPEIAANHITQAVLT 120
 QY 124 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 183
 Db 121 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 180
 QY 184 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
 Db 181 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
 QY 244 GGNPADEEAAQWGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303
 Db 241 GGNPADEEAAQWGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300
 QY 304 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGSTRPGLVAPAPLAQEREDDED 363
 Db 301 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGSTRPGLVAPAPLAQEREDDED 360
 QY 364 DWDEEDDW 371
 Db 361 DWDEEDDW 368

RESULT 5

US-09-072-596-109
 ; Sequence 109, Application US/09072596
 ; Patent No. 6458366
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Campos-Neto, Antonia
 ; APPLICANT: Houghton, Raymond
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Twardzik, Daniel R.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Hendrickson, Ronald C.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
 ; TUBERCULOSIS
 ; NUMBER OF SEQUENCES: 350
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED AND BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 05-MAY-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.417C9
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 109:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 368 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-09-072-596-109

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MLWHPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60

QY 64 SDKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

QY 124 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180

QY 184 STTNPIFGMPSPGSSTPVQQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSPGSSTPVQQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 240

QY 244 GGNPADEAAQMGLLGTSPLSNHPPLAGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303
DB 241 GGNPADEAAQMGLLGTSPLSNHPPLAGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300

QY 304 EKPVPSPVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 363
DB 301 EKPVPSPVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 360

QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368

RESULT 6

US-09-072-967-114
; Sequence 114, Application US/09072967
; Patent No. 6592877
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-072-967-114

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MLWHPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60

QY 64 SDKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

QY 124 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180

QY 184 STTNPIFGMPSPGSSTPVQQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSPGSSTPVQQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 240

QY 244 GGNPADEAAQMGLLGTSPLSNHPPLAGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303
DB 241 GGNPADEAAQMGLLGTSPLSNHPPLAGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300

QY 304 EKPVPSPVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 363
DB 301 EKPVPSPVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 360

QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368

RESULT 7

US-10-193-002-109
; Sequence 109, Application US/10193002
; Patent No. 6949246
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TITLE OF INVENTION: TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:

APPLICATION NUMBER: US/10/193,002
FILING DATE: 10-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-10-193-002-109

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTOAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTOAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTENDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTPIALTENDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSPGSSSTPVQQLPPAATQTTLGQMGSGPMQOQTQPLQOQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSPGSSSTPVQQLPPAATQTTLGQMGSGPMQOQTQPLQOQVTSLSFQVGGTG 240
QY 244 GGNPADEEAQMGLLGTSPLSNHLPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 303
DB 241 GGNPADEEAQMGLLGTSPLSNHLPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 300
QY 304 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 363
DB 301 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 360
QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368

RESULT 8
US-10-084-843-114
Sequence 114, Application US/10084843
Patent No. 6962710
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 114:
US-10-084-843-114

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTOAMATTPSLPEIAANHITQAVLT 123
DB 61 SKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTOAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTENDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTPIALTENDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSPGSSSTPVQQLPPAATQTTLGQMGSGPMQOQTQPLQOQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSPGSSSTPVQQLPPAATQTTLGQMGSGPMQOQTQPLQOQVTSLSFQVGGTG 240
QY 244 GGNPADEEAQMGLLGTSPLSNHLPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 303
DB 241 GGNPADEEAQMGLLGTSPLSNHLPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 300
QY 304 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 363
DB 301 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 360
QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368

RESULT 9
US-09-073-009-126
Sequence 126, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark


```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-09-287-849-2

Query Match      17.5%; Score 330.5; DB 2; Length 729;
Best Local Similarity 28.9%; Pred. No. 6.8e-20;
Matches 120; Conservative 45; Mismatches 173; Indels 77; Gaps 13;

QY      8  AMPPELNTARLMAGAGPAPMLAAAAGWQTL-----SAALDAQAVELTARLNSLGEAWTGG 62
Db      147  ALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVMGLTVGS-----WIGS 202

QY      63  GSDKALAAATPMVWLQTAQKTRAMQATAQAAAAYTOAMATTTPSLPEIAANHITQAVL 122
Db      203  SAGLMVAASPVMVMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMIL 262

QY      123  TATNFFGINTIPALTENDYFIRMNQALAMEVVOAETA VNT----- 165
Db      263  IATNLGQNTPAIAVNEAEYG-EMWQDAAAAMFGYAAATATATATATLLPPEEAPEMTSAGG 321

QY      166  LFEKLEPMASILDPCASQSTTNPI-----FCMPSPGSGSTPVGQL-----PPAATQTL 212
Db      322  LLEQAAAVEEASDTAAANQLMNNVPQALQQLAQPTQG-TTPSSKLGGLWKTVPSPHRSPIS 380

QY      213  GOLGEMSGPMQOLTQPLQOVTSLFSQVGG--TGGGNPADEEAAQWGLLGTSPLSNHPLAG 270
Db      381  NMVSWANNHMSMTNSGVSMNTLSLWLGFPAPAAARQAVQTAAGVRAVMSLGLSSLSGSS 440

QY      271  GSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLIEKPVP-----SVMPPAAAAGSSAT 322
Db      441  GLGGGVAANLGRAASV-----GSLS-VPQAWAANAQAVTPAARALPLTSLTSAERGPQM 495

QY      323  GGAAPVAGAMGQGAQSGG-----STRPGLVAPAPLAQREED 360
Db      496  LGGLPVQG----MGARAGGGLSGVLRVPPRPYVMPHSPAAGDIAPPALSQDRFAD 546
```

```
RESULT 14
US-09-287-849-26
; Sequence 26, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
```

```
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
US-09-287-849-26
```

```
Query Match      17.4%; Score 328.5; DB 2; Length 596;
Best Local Similarity 28.9%; Pred. No. 7.7e-20;
Matches 120; Conservative 45; Mismatches 173; Indels 77; Gaps 13;

QY      8  AMPPELNTARLMAGAGPAPMLAAAAGWQTL-----SAALDAQAVELTARLNSLGEAWTGG 62
Db      14  ALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVMGLTVGS-----WIGS 69

QY      63  GSDKALAAATPMVWLQTAQKTRAMQATAQAAAAYTOAMATTTPSLPEIAANHITQAVL 122
Db      70  SAGLMVAASPVMVMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMIL 129

QY      123  TATNFFGINTIPALTENDYFIRMNQALAMEVVOAETA VNT----- 165
Db      130  IATNLGQNTPAIAVNEAEYG-EMWQDAAAAMFGYAAATATATATATLLPPEEAPEMTSAGG 188

QY      166  LFEKLEPMASILDPCASQSTTNPI-----FCMPSPGSGSTPVGQL-----PPAATQTL 212
Db      189  LLEQAAAVEEASDTAAANQLMNNVPQALQQLAQPTQG-TTPSSKLGGLWKTVPSPHRSPIS 247

QY      213  GOLGEMSGPMQOLTQPLQOVTSLFSQVGG--TGGGNPADEEAAQWGLLGTSPLSNHPLAG 270
Db      248  NMVSWANNHMSMTNSGVSMNTLSLWLGFPAPAAARQAVQTAAGVRAVMSLGLSSLSGSS 307

QY      271  GSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLIEKPVP-----SVMPPAAAAGSSAT 322
Db      308  GLGGGVAANLGRAASV-----GSLS-VPQAWAANAQAVTPAARALPLTSLTSAERGPQM 362

QY      323  GGAAPVAGAMGQGAQSGG-----STRPGLVAPAPLAQREED 360
Db      363  LGGLPVQG----MGARAGGGLSGVLRVPPRPYVMPHSPAAGDIAPPALSQDRFAD 413
```

```
RESULT 15
US-09-073-009-142
; Sequence 142, Application US/09073009
; Patent No. 6555653
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,009
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
```

TUBERCULOSIS AND ME

REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-009-142

Query Match 17.2%; Score 326; DB 2; Length 423;
Best Local Similarity 30.4%; Pred. No. 7.9e-20;
Matches 121; Conservative 45; Mismatches 164; Indels 68; Gaps 17;

QY	9	MPELNTARLMAGAGPAPMLAAAGWQTLISAALDAQAVELTARLNSL-GEAWTGGSDKA	67
Db	6	LPPEVNSRMVSGPGPEKSLAAAANDGVAAELTSAAYSGVSVTLIVEPMWGPAAAAAM	65
QY	68	LAATPMVVLQSTASTOAKTRAMQATAQAAYTQAMATT--PSLPEIAANHITQAVLTAT	125
Db	66	AAATPYVGLAATRAALAKETATQAAAAEAFTAFAMTVPESL--VAANRSLMSLVAA	123
QY	126	NFFGINTIPIALTEMDFIRMNQALAMEVYQAEAVNTLFEKLEPM---ASILDPGAS	182
Db	124	NILQNSAAIATQAEY-AEMWAQDAAVMYSYEGASAAASALPPFTPPVQGTGPAGPAAA	182
QY	183	QSTTNPITGMPSPGSGSTPVGQLPP-----AATQTLGOLGEMS--GPMQQLTQ	227
Db	183	AAATQAAGAGAVADAQATLAQLPPGILSDILSALAANDPLTSLGLGIATSTLNPOVGSQAQ	242
QY	228	PL-----QQVTSLSFQVGGTGGGNPADEEAAQ---MGLLGTG---PLSNHPLA---	269
Db	243	PIVITPIGELDVIALYIASIATGSIATATNTARPHWHLGYGNAGGLGPTQGHPLSSAT	302
QY	270	-----GGSGP-SAGAG---LLRAESLPGAGGSLTRTPIMSQLIEKVPVPSVMPAA	315
Db	303	DEPEPHWGPFGGAAPVSGVGHAAVLGALSVPHSW--TTAAPEIQIAVQ--ATPTFFSSA	358
QY	316	AGSSATGCAAPVG-----AGAMGQGAQSGGSTRPG	346
Db	359	GADPTALNG-MPAGLLSGMALASLAARGTTGGGTRSG	395

Search completed: January 5, 2006, 21:19:28
Job time : 47 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2006, 16:38:03 ; Search time 138 Seconds

(without alignments)
1181.228 Million cell updates/sec

Title: US-09-872-505A-6

Perfect score: 1892

Sequence: 1 MITMLWHAMPFELNTARLMA.....PLAQEREDDDWDEEDDW 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq 21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1892	100.0	371	AAW72929	Mycobacte
2	1892	100.0	371	AAW72929	Mycobacte
3	1878	99.3	368	AAW19844	Mycobacte
4	1878	99.3	368	AAW19844	Mycobacte
5	1873	99.0	368	AAW32384	Mycobacte
6	1873	99.0	368	AAW32384	Mycobacte
7	1873	99.0	368	AAW32384	Mycobacte
8	1873	99.0	368	AAW32384	Mycobacte
9	1873	99.0	368	AAW32384	Mycobacte
10	1873	99.0	368	AAW32384	Mycobacte
11	1873	99.0	368	AAW32384	Mycobacte
12	1873	99.0	368	AAW32384	Mycobacte
13	1873	99.0	368	AAW32384	Mycobacte
14	1873	99.0	368	AAW32384	Mycobacte
15	1873	99.0	368	AAW32384	Mycobacte
16	1873	99.0	368	AAW32384	Mycobacte
17	1873	99.0	368	AAW32384	Mycobacte
18	1873	99.0	368	AAW32384	Mycobacte
19	1873	99.0	368	AAW32384	Mycobacte
20	1873	99.0	368	AAW32384	Mycobacte
21	1873	99.0	368	AAW32384	Mycobacte
22	1873	99.0	368	AAW32384	Mycobacte
23	1873	99.0	368	AAW32384	Mycobacte
24	1873	99.0	368	AAW32384	Mycobacte

25	333.5	17.6	729	7	ADA26373	Ada26373	Mycobacte
26	330.5	17.5	729	5	AAE29708	AAE29708	Mycobacte
27	330.5	17.5	729	5	AAE17572	AAE17572	Mycobacte
28	330.5	17.5	729	7	ADF69753	ADF69753	Fusion pr
29	330.5	17.5	729	7	ADA26356	ADA26356	Mycobacte
30	330	17.4	1154	7	ADA26368	ADA26368	Mycobacte
31	328.5	17.4	596	2	AAV32070	AAV32070	Mycobacte
32	328.5	17.4	596	5	AAE29710	AAE29710	Mycobacte
33	328.5	17.4	596	5	AAE17574	AAE17574	Mycobacte
34	328.5	17.4	596	7	ADF69777	ADF69777	Fusion pr
35	328.5	17.4	599	5	AAU74599	AAU74599	Antigenic
36	328.5	17.4	729	4	AAO22142	AAO22142	Ral12-H9-3
37	328.5	17.4	729	5	AAE29709	AAE29709	Mycobacte
38	328.5	17.4	729	5	AAE17573	AAE17573	Mycobacte
39	328.5	17.4	729	7	ADA26374	ADA26374	Mycobacte
40	328.5	17.4	813	7	ADA26367	ADA26367	Mycobacte
41	328.5	17.4	825	7	ADA26366	ADA26366	Mycobacte
42	328.5	17.4	875	7	ADA26365	ADA26365	Mycobacte
43	328.5	17.4	930	5	AAE29731	AAE29731	Mycobacte
44	328.5	17.4	930	7	ADA26364	ADA26364	Mycobacte
45	328.5	17.4	1016	7	ADA26370	ADA26370	M. bovis

ALIGNMENTS

RESULT 1	
AAW72929	
ID	AAW72929 standard; protein; 371 AA.
XX	
AC	AAW72929;
XX	
DT	21-JAN-1999 (first entry)
XX	
DE	Mycobacterium tuberculosis antigen RD1-ORF5.
XX	
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
KW	infection.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	W09844119-A1.
XX	
PD	08-OCT-1998.
XX	
PF	01-APR-1998; 98WO-DK000132.
XX	
PR	02-APR-1997; 97DK-00000376.
PR	18-APR-1997; 97US-0044624P.
PR	10-NOV-1997; 97DK-00001277.
PR	05-JAN-1998; 98US-0070488P.
XX	
PA	(STAT-) STATENS SERUM INST.
XX	
PI	Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB;
PI	Oettinger T, Florio W;
XX	
XX	WPI; 1998-542705/46.
DR	N-PSDB; AAV63939.
DR	
XX	
PT	New isolated mycobacteria polypeptides and nucleic acids - used for
PT	developing products for the diagnosis of or vaccination against
PT	mycobacterial infections, particularly tuberculosis.
XX	
PS	Claim 1; Page 200-202; 163pp; English.
XX	
CC	The present sequence represents a Mycobacterium tuberculosis protein.
CC	Products from the present invention, which describes protein fragments
CC	and nucleic acid fragments derived from M.tuberculosis, can be used in
CC	the detection of and prevention of mycobacterial infections. In
CC	particular, the proteins and nucleic acids can be used for the diagnosis
CC	of or vaccination against tuberculosis caused by M. tuberculosis, M.
CC	africanum or M. bovis

```

XX SQ Sequence 371 AA;
Query Match 100.0%; Score 1892; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLTSAALDAQAVELTARLNSLGEAWT 60
DB 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLTSAALDAQAVELTARLNSLGEAWT 60

QY 61 GGGSDKALAAATPMVWLQTAQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQA 120
DB 61 GGGSDKALAAATPMVWLQTAQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQA 120

QY 121 VLTATNPFGINPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPG 180
DB 121 VLTATNPFGINPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPG 180

QY 181 ASQSTTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVG 240
DB 181 ASQSTTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVG 240

QY 241 GTGGGNPADEEAQMGLLGTSLPSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300
DB 241 GTGGGNPADEEAQMGLLGTSLPSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300

QY 301 QLIEKPVAPSVMPAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360
DB 301 QLIEKPVAPSVMPAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360

QY 361 DEDDWDDEDDW 371
DB 361 DEDDWDDEDDW 371

RESULT 2
AAY21946
ID AAY21946 standard; protein; 371 AA.
XX
AC AAY21946;
XX
DT 06-SEP-1999 (first entry)
XX
DE Amino acid sequence of antigen RD1-ORF5.
XX
KW Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KW CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
KW CFP25A; CFP30B; CFP7B.
XX
OS Mycobacterium tuberculosis.
XX
PN W09924577-A1.
XX
PD 20-MAY-1999.
XX
PF 08-OCT-1998; 98WO-DK000438.
XX
PR 10-NOV-1997; 97DK-00001277.
PR 05-JAN-1998; 98US-0070488P.
PR 01-APR-1998; 98WO-DK000132.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Andersen P, Skjot R;
XX
XX WPI; 1999-347282/29.
DR N-PSDB; AAX81046.
XX
XX New immunogenic fragment of Mycobacterium tuberculosis.
XX

```

```

PS Example 2; Page 219-220; 265pp; English.
XX
CC The invention describes a substantially pure immunogenic polypeptide
CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC protective immune response against infections by mycobacteria belonging
CC to the tuberculosis complex. The invention provides a (1) fusion
CC polypeptide comprising at least one polypeptide fragment (I) and at least
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
CC different amino acid sequence from M. tuberculosis, and/or including a
CC sequence which protects the first amino acid sequence from in vivo
CC degradation or post-translational processing; (3) a nucleic acid fragment
CC that encodes the above polypeptides. The polypeptides and nucleic acid
CC are useful as pharmaceuticals, for diagnosis of and as antigens for
CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC bovis. The polypeptides are also useful for diagnosing ongoing or
CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CFP7A or
CC CFP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine
XX
SQ Sequence 371 AA;
Query Match 100.0%; Score 1892; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLTSAALDAQAVELTARLNSLGEAWT 60
DB 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLTSAALDAQAVELTARLNSLGEAWT 60

QY 61 GGGSDKALAAATPMVWLQTAQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQA 120
DB 61 GGGSDKALAAATPMVWLQTAQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQA 120

QY 121 VLTATNPFGINPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPG 180
DB 121 VLTATNPFGINPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPG 180

QY 181 ASQSTTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVG 240
DB 181 ASQSTTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQQLTQPLQVTSLSFQVG 240

QY 241 GTGGGNPADEEAQMGLLGTSLPSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300
DB 241 GTGGGNPADEEAQMGLLGTSLPSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300

QY 301 QLIEKPVAPSVMPAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360
DB 301 QLIEKPVAPSVMPAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360

QY 361 DEDDWDDEDDW 371
DB 361 DEDDWDDEDDW 371

RESULT 3
AAB19844
ID AAB19844 standard; protein; 368 AA.
XX
AC AAB19844;
XX
XX 05-MAR-2001 (first entry)
DT
DE Mycobacterium tuberculosis protein MTB3.
XX
XX MTB3; tuberculosis; BCG; vaccine; infection; diagnosis.
XX

```

OS Mycobacterium tuberculosis.
 XX WO200066157-A1.
 XX 09-NOV-2000.
 XX 04-MAY-2000; 2000WO-US012257.
 XX 04-MAY-1999; 99US-0132505P.
 XX (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
 XX PA Gennaro ML;
 XX PI WPI; 2001-007153/01.
 XX DR N-PSDB; AAA89037.
 XX PT Novel polypeptide encoded by open reading frames present in Mycobacterium
 XX tuberculosis genome and not by the BCG strain of M. bovis, useful as
 XX vaccine and for diagnosing tuberculosis infection.
 XX PS Claim 11; Fig 1; 35pp; English.
 XX CC The present sequence is that of the Mycobacterium tuberculosis MTBN3
 XX protein. This is 1 of 8 proteins, i.e. MTBN1-8 (see AAB19842-49), encoded
 XX by 8 open reading frames (see AAA89035-42) identified as being present in
 XX the genome of M. tuberculosis but absent from the genome of the BCG
 XX strain of Mycobacterium bovis. MTBN1-8 represent reagents that are useful
 XX in discriminating between M. tuberculosis and BCG and, in particular, for
 XX diagnostic methods which discriminate between exposure of a subject to M.
 XX tuberculosis and vaccination with BCG. The invention features these MTBN
 XX polypeptides, functional fragments of them, DNA encoding them, vectors,
 XX transformed cells, and diagnostic, therapeutic, and prophylactic
 XX (vaccine) methods, including genetic vaccination methods
 XX SQ Sequence 368 AA;
 Query Match 99.3%; Score 1878; DB 4; Length 368;
 Best Local Similarity 100.0%; Pred. No. 2.3e-130;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 MLWHAMPPEINTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
 DB 1 MLWHAMPPEINTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
 QY 64 SDKALAAATPMVVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
 DB 61 SDKALAAATPMVVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
 QY 124 ATNFFGINTPIALTENDYFIRMNQALAMEYQAEAVNTLFEKLEPMASILDPCASQ 183
 DB 121 ATNFFGINTPIALTENDYFIRMNQALAMEYQAEAVNTLFEKLEPMASILDPCASQ 180
 QY 184 STTNPIFGMPSPGSGSTPVGQLPPAAATQTLGOLGEMSGPMQOLTQPLQOQVTSLSFQVGGTG 243
 DB 181 STTNPIFGMPSPGSGSTPVGQLPPAAATQTLGOLGEMSGPMQOLTQPLQOQVTSLSFQVGGTG 240
 QY 244 GGNPADEEAAQMGILLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 303
 DB 241 GGNPADEEAAQMGILLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 300
 QY 304 EKPVPSPVMPAAAGSATGGAAPVGAAGMGQAGGSGTRPGLVAPAPLAQREDEDED 363
 DB 301 EKPVPSPVMPAAAGSATGGAAPVGAAGMGQAGGSGTRPGLVAPAPLAQREDEDED 360
 QY 364 DWDEEDDW 371
 DB 361 DWDEEDDW 368
 RESULT 4
 ADG15733
 ID ADG15733 standard; protein; 368 AA.

XX ADG15733;
 XX AC 26-FEB-2004 (first entry)
 XX DT M tuberculosis Rv3873 protein for generating peptide epitopes.
 XX DE
 XX KW antibacterial; antitubercular; tuberculostatic; vaccine;
 XX KW diagnostic reagent; epitope; RD1 region; RD2 region;
 XX KW Mycobacterium tuberculosis; Mycobacterium bovis; Mycobacterium africanum.
 XX OS Mycobacterium tuberculosis.
 XX XX WO2003093307-A2.
 XX PD 13-NOV-2003.
 XX XX 28-APR-2003; 2003WO-GB001815.
 XX PF 27-APR-2002; 2002GB-00009723.
 XX PR 27-APR-2002; 2002GB-00009724.
 XX PS (UKEN-) UK SEC FOR ENVIRONMENT FOOD & RURAL AFF.
 XX PA Cockle PJ, Vordermeier HM, Gordon SV, Hewinson RG;
 XX PI WPI; 2003-903652/82.
 XX XX New diagnostic reagents comprising a peptide having an epitope from
 XX PT polypeptides Rv1986, Rv3878, Rv1983, Rv3873 or Rv3879 derived from RD1
 XX and RD2 regions of Mycobacterium, useful as vaccines against
 XX PT Mycobacterium infections.
 XX XX Claim 19; SEQ ID NO 5; 77pp; English.
 XX CC The invention relates to a novel diagnostic reagent comprising a peptide
 XX epitope from a protein encoded by the RD1 or RD2 regions of the
 XX Mycobacterium tuberculosis, M. bovis or M. africanum genome. The
 XX diagnostic reagents comprised within the kit are selected so that they
 XX are able to differentiate between M. bovis, M. tuberculosis or M.
 XX africanum-infected mammals and mammals vaccinated against M. bovis, M.
 XX tuberculosis or M. africanum. The polypeptide or its variant or fragment
 XX is useful as a medicament provided that the polypeptide is not a MPT-64
 XX polypeptide or a polypeptide encoded by the Rv1984C region of the M.
 XX bovis, M. tuberculosis or M. africanum genomes. The nucleic acid which
 XX encodes the polypeptide or the diagnostic reagents can be used as a
 XX vaccine, especially against Mycobacterium infections. This sequence
 XX corresponds to a protein used to derive the peptide epitopes used in the
 XX method of the invention.
 XX SQ Sequence 368 AA;
 Query Match 99.3%; Score 1878; DB 7; Length 368;
 Best Local Similarity 100.0%; Pred. No. 2.3e-130;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 MLWHAMPPEINTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
 DB 1 MLWHAMPPEINTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
 QY 64 SDKALAAATPMVVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
 DB 61 SDKALAAATPMVVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
 QY 124 ATNFFGINTPIALTENDYFIRMNQALAMEYQAEAVNTLFEKLEPMASILDPCASQ 183
 DB 121 ATNFFGINTPIALTENDYFIRMNQALAMEYQAEAVNTLFEKLEPMASILDPCASQ 180
 QY 184 STTNPIFGMPSPGSGSTPVGQLPPAAATQTLGOLGEMSGPMQOLTQPLQOQVTSLSFQVGGTG 243
 DB 181 STTNPIFGMPSPGSGSTPVGQLPPAAATQTLGOLGEMSGPMQOLTQPLQOQVTSLSFQVGGTG 240
 QY 244 GGNPADEEAAQMGILLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 303

```
||||| 241 GGNPADEAAQWGLLGTSPLSNHPFLAGSGSPSAGAGLLRAESLPGAGSLTRTPMSQLI 300
||||| 304 EKPVPASVMPAAAAAGSSATGGGAAPVGAGAMGGAGGSGSTRFGLVAPAPLAQEREDEDED 363
||||| 301 EKPVPASVMPAAAAAGSSATGGGAAPVGAGAMGGAGGSGSTRFGLVAPAPLAQEREDEDED 360
||||| 364 DWDEDDW 371
||||| 361 DWDEDDW 368

RESULT 5
AAW32384
ID AAW32384 standard; protein; 368 AA.
AC AAW32384;
DT 13-JAN-1998 (first entry)
DE Mycobacterium tuberculosis antigen Tb37-FL.
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
OS Mycobacterium tuberculosis.
FH Key Location/Qualifiers
FT Misc-difference 10 /note= "Any amino acid"
FN WO9709429-A2.
PD 13-MAR-1997.
PF 30-AUG-1996; 96WO-US014675.
PR 01-SEP-1995; 95US-00523435.
PR 22-SEP-1995; 95US-00532136.
PR 22-MAR-1996; 96US-00620280.
PR 05-JUN-1996; 96US-00658800.
PR 12-JUL-1996; 96US-00680573.
PA (CORI-) CORIXA CORP.
XX Read SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX WPI; 1997-192904/17.
XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
FT useful for diagnosis of M. tuberculosis infection.
XX Example 3; Page 159-161; 190pp; English.
CC A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC Tb37-FL. The immunogenic polypeptide can be used to diagnose
CC M.tuberculosis infection by forming complexes with specific antibodies in
CC the sample. Fragments of DNA encoding the immunogenic polypeptide can be
CC used as diagnostic primers or probes and agents that bind to the antigen,
CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
CC also used for diagnosis
SQ Sequence 368 AA;
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELTARLMAGAGPAPMLAAAGWQTLSAALDAQVELTARLNSLGEAWTGGG 63
```

```
||||| 1 MLWHAMPPEXTARLMAGAGPAPMLAAAGWQTLSAALDAQVELTARLNSLGEAWTGGG 60
||||| 64 SDKALAAATPMVWLQTAQTAKTRAMQATAQAAAAYTQAMATTPLSPPIAANHITQAVLT 123
||||| 61 SDKALAAATPMVWLQTAQTAKTRAMQATAQAAAAYTQAMATTPLSPPIAANHITQAVLT 120
||||| 124 ATNPFGINITIPALTENDYFIRMKNQALANEVYQAEAVNTLREKLEPMASILLDPGASQ 183
||||| 121 ATNPFGINITIPALTENDYFIRMKNQALANEVYQAEAVNTLREKLEPMASILLDPGASQ 180
||||| 184 STTNPIFGMPSPGSTPVGQLPPAATOTLGLGEMSGPMQQLTOPLQOVTSLFSQVGGTG 243
||||| 181 STTNPIFGMPSPGSTPVGQLPPAATOTLGLGEMSGPMQQLTOPLQOVTSLFSQVGGTG 240
||||| 244 GGNPADEAAQWGLLGTSPLSNHPFLAGSGSPSAGAGLLRAESLPGAGSLTRTPMSQLI 303
||||| 241 GGNPADEAAQWGLLGTSPLSNHPFLAGSGSPSAGAGLLRAESLPGAGSLTRTPMSQLI 300
||||| 304 EKPVPASVMPAAAAAGSSATGGGAAPVGAGAMGGAGGSGSTRFGLVAPAPLAQEREDEDED 363
||||| 301 EKPVPASVMPAAAAAGSSATGGGAAPVGAGAMGGAGGSGSTRFGLVAPAPLAQEREDEDED 360
||||| 364 DWDEDDW 371
||||| 361 DWDEDDW 368

RESULT 6
AAW32452
ID AAW32452 standard; protein; 368 AA.
AC AAW32452;
DT 09-JAN-1998 (first entry)
DE Mycobacterium tuberculosis antigen Tb37-FL.
XX Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
OS Mycobacterium tuberculosis.
FH Key Location/Qualifiers
FT Misc-difference 10 /note= "Any amino acid"
FN WO9709428-A2.
PD 13-MAR-1997.
PF 30-AUG-1996; 96WO-US014674.
PR 01-SEP-1995; 95US-00523436.
PR 22-SEP-1995; 95US-00533634.
PR 22-MAR-1996; 96US-00620874.
PR 05-JUN-1996; 96US-00659683.
PR 12-JUL-1996; 96US-00680574.
PA (CORI-) CORIXA CORP.
XX Read SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX WPI; 1997-192903/17.
XX New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
FT useful in vaccines for prevention or treatment of tuberculosis, also for
FT diagnosis.
XX Example 3; Page 146-147; 168pp; English.
XX A new immunogenic polypeptide has been developed comprising an
```

CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC Tb37-FL The immunogenic protein, and fusion proteins containing one or
CC more of the proteins or one of the proteins plus ESAT-6, are useful in
CC vaccines, preferably when formulated with a non-specific adjuvant, to
CC induce an immune response against M.tuberculosis (for treatment or
CC prevention)
XX
XX Sequence 368 AA;

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLNDAQAQVELTARLNSLGEAWTGGG 63
Db 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWQTLNDAQAQVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
Db 61 SDKALAAATPMVWVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
Db 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSPGSSTPVQQLPPAATQTQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
Db 181 STTNPIFGMPSPGSSTPVQQLPPAATQTQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
QY 244 GGNPADEEAAQMGLLGTSPLSNHPLAGSGSPSAGALLRAESLPGAGGSLTRTPMSQLI 303
Db 241 GGNPADEEAAQMGLLGTSPLSNHPLAGSGSPSAGALLRAESLPGAGGSLTRTPMSQLI 300
QY 304 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDED 363
Db 301 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDED 360
QY 364 DWDEEDDW 371
Db 361 DWDEEDDW 368

RESULT 7
AAW64338
ID AAW64338 standard; protein; 368 AA.

XX AAW64338;
XX
XX 17-OCT-2003 (revised)
DT 09-NOV-1998 (first entry)
XX Mycobacterium tuberculosis antigen Tb37-FL.
XX Tuberculosis; infection; diagnosis; antigen; Tb37-FL.
XX Mycobacterium tuberculosis; strain H37Rv.

Key Location/Qualifiers
FT Misc-difference 10 /note= "unidentified"

XX WO9816645-A2.
XX 23-APR-1998.
XX 07-OCT-1997; 97WO-US018214.
XX 11-OCT-1996; 96US-00729622.
XX 13-MAR-1997; 97US-00818111.
XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedwick TS, Twardzik DR, Lodes MJ;
XX WPI; 1998-251292/22.
XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
PT develop products for the detection of M. tuberculosis infection and
PT diagnosis of tuberculosis.
XX
XX Example 3; Page 143-144; 250pp; English.

XX This polypeptide comprises Mycobacterium tuberculosis antigen Tb37-FL. It
CC is encoded by genomic DNA isolated from a M. tuberculosis strain H37Rv
CC genomic library using a probe from clone Tb38-1 (see AAV44384). The
CC invention relates to compositions and methods for diagnosing
CC tuberculosis. It provides polypeptides (see AAW64291-W64379) comprising
CC an antigenic portion of a soluble M. tuberculosis antigen, or an
CC immunogenic portion of an M. tuberculosis antigen, as well as DNA
CC sequences encoding such polypeptides, recombinant expression vectors and
CC transformed or transfected host cells. Also claimed are methods and
CC diagnostic kits for detecting M. tuberculosis infection in a patient
CC using these polypeptides, antibodies or oligonucleotide probes and
CC primers, for the diagnosis of tuberculosis. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX Sequence 368 AA;

Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLNDAQAQVELTARLNSLGEAWTGGG 63
Db 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWQTLNDAQAQVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
Db 61 SDKALAAATPMVWVLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
Db 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSPGSSTPVQQLPPAATQTQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
Db 181 STTNPIFGMPSPGSSTPVQQLPPAATQTQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
QY 244 GGNPADEEAAQMGLLGTSPLSNHPLAGSGSPSAGALLRAESLPGAGGSLTRTPMSQLI 303
Db 241 GGNPADEEAAQMGLLGTSPLSNHPLAGSGSPSAGALLRAESLPGAGGSLTRTPMSQLI 300
QY 304 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDED 363
Db 301 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDED 360
QY 364 DWDEEDDW 371
Db 361 DWDEEDDW 368

RESULT 8
AAW81705
ID AAW81705 standard; protein; 368 AA.

XX AAW81705;
XX
XX 27-JAN-1999 (first entry)
XX M. tuberculosis immunogenic polypeptide Tb37-FL.
XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
XX vaccine; pharmaceutical; infection; diagnosis.
KW

```
XX OS Mycobacterium tuberculosis.
XX FH Key Location/Qualifiers
XX FT Misc-difference 10
XX FT /label= unknown
XX FT
XX PN WO9816646-A2.
XX PD 23-APR-1998.
XX XX
XX PF 07-OCT-1997; 97WO-US018293.
XX XX
XX PR 11-OCT-1996; 96US-00730510.
XX PR 13-MAR-1997; 97US-00818112.
XX XX
XX PA (CORI-) CORIXA CORP.
XX PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedwick TS, Twardzik DR, Lodes MJ;
XX DR WPI; 1998-261042/23.
XX XX
XX PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
XX FT develop products for the detection of M. tuberculosis infection and for
XX FT diagnosis, treatment and prevention of tuberculosis.
XX XX
XX PS Example 3B; Page 137-138; 230pp; English.
XX XX
XX CC This sequence represents an immunogenic portion of a soluble
XX CC Mycobacterium tuberculosis (MT) antigen which can be used in a method for
XX CC inducing protective immunity against tuberculosis (TB). This sequence can
XX CC be formulated into vaccines and/or pharmaceutical compositions for
XX CC immunising against M. tuberculosis infection or may be used for the
XX CC diagnosis of tuberculosis
XX XX
XX SQ Sequence 368 AA;
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAAATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAAATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSGSPVQQLPPAATQTLGQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSGSPVQQLPPAATQTLGQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQMLLGTSPLSNHPHLAGGSPSAGAGLLRAESLPGAGGSLTRTPMLMSQLI 303
DB 241 GGNPADEAAQMLLGTSPLSNHPHLAGGSPSAGAGLLRAESLPGAGGSLTRTPMLMSQLI 300
QY 304 EKPVAPSYMPPAAAAGSSATGGAAPVGAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 363
DB 301 EKPVAPSYMPPAAAAGSSATGGAAPVGAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 360
QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368
```

```
ID AAY38992 standard; protein; 368 AA.
XX AC AAY38992;
XX DT 05-NOV-1999 (first entry)
XX DE M. tuberculosis recombinant antigen protein Tb37-Fl.
XX KW Antigen; diagnosis; detection; infection; antibody; immunisation;
XX KW vaccine; immunity.
XX OS Mycobacterium tuberculosis.
XX PN WO9942118-A2.
XX PD 26-AUG-1999.
XX PF 17-FEB-1999; 99WO-US003265.
XX PR 18-FEB-1998; 98US-00024753.
XX PR 05-MAY-1998; 98US-00072596.
XX PA (CORI-) CORIXA CORP.
XX PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedwick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX DR WPI; 1999-527416/44.
XX XX
XX PT New polypeptide comprising antigenic portions of M. tuberculosis.
XX PS Example 3; Page 177-179; 323pp; English.
XX CC This invention describes novel recombinant antigens and their encoding
XX CC nucleic acids derived from Mycobacterium tuberculosis. The novel
XX CC polypeptides are useful for detecting M. tuberculosis infection in a
XX CC biological sample by detecting antibodies which bind with the
XX CC polypeptides, and are useful as vaccines for immunizing against M.
XX CC tuberculosis infection. The new detection methods are needed as current
XX CC vaccination strategies do not provide 100% immunity
XX SQ Sequence 368 AA;
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAAATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAAATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTIPALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSGSPVQQLPPAATQTLGQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSGSPVQQLPPAATQTLGQLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQMLLGTSPLSNHPHLAGGSPSAGAGLLRAESLPGAGGSLTRTPMLMSQLI 303
DB 241 GGNPADEAAQMLLGTSPLSNHPHLAGGSPSAGAGLLRAESLPGAGGSLTRTPMLMSQLI 300
QY 304 EKPVAPSYMPPAAAAGSSATGGAAPVGAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 363
DB 301 EKPVAPSYMPPAAAAGSSATGGAAPVGAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 360
QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368
```


39	IAAKPMITWLSQA-----AEQTTQAEBAHQRMASSTPGMAVITENHHTQAILATINP	91
128	FGINTIPIALTMDYFIRIMWQAALAMVEVYQAEAVNTLFEKLEPMASILDPGA-----	181
92	FGINMAPIAFTEAGDFICWRTQTALAMNSYQAEITLLNTAFQKLEPMAAILNPSSPPSA	151
182	--SQ--STTNPIFGHPSFGSSTPVQOLPPAATOTLQQLGEMSGPMQOLTOPLQVYTSLS	237
152	LTSQVNQFTQMISGFSAAIPSTQVLQ-----QTVGVQVAE-----LAAPMQQVKSLEF	198
238	QVGGTGGGNPAD----EBAAGMLGTTGPTLNHPPLAGSGSPSAGALLRAESLPGAGGSL	293
199	SIDSTGVYTSAQRGDTESNHRIGLFGASTLSHPVLGVTGTTTTRLLCAESLPSAGSL	258
294	TRTPPLMS--QLIEKPVPAS-----VMPAAAGSSSATGGAAPVGAGAGQAQSGGST	343
259	AWTFLMTOPOLDKSIGIAPRQVNLV-----PWAAGSPGHNAQDGGTT	302

RESULT 12	
ADGI5750	
ID	ADGI5750 standard; protein; 100 AA.
XX	
XX	
AC	ADGI5750;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	RV3873 sequence fragment used to generate epitopes.
XX	
XX	
KW	antibacterial; antitubercular; tuberculostatic; vaccine;
KW	diagnostic reagent; epitope; RD1 region; RD2 region;
KW	Mycobacterium tuberculosis; Mycobacterium bovis; Mycobacterium africanum.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	WO2003093307-A2.
XX	
PD	13-NOV-2003.
XX	
PF	28-APR-2003; 2003WO-GB001815.
XX	
PR	27-APR-2002; 2002GB-00009723.
PR	27-APR-2002; 2002GB-00009724.
XX	
PA	(UKEN-) UK SEC FOR ENVIRONMENT FOOD & RURAL AFF.
XX	
PI	Cockle PJ, Vordermeier HM, Gordon SV, Hewinson RG;
XX	
DR	WPI; 2003-903652/82.

XX New diagnostic reagents comprising a peptide having an epitope from
PT polypeptides Rv1986, Rv3878, Rv1983, Rv3873 or Rv3879 derived from RD1
PT and RD2 regions of Mycobacterium, useful as vaccines against
PT Mycobacterium infections.
XX
PS Disclosure; SEQ ID NO 23; 77pp; English.
XX
XX The invention relates to a novel diagnostic reagent comprising a peptide
XX epitope from a protein encoded by the RD1 or RD2 regions of the
CC Mycobacterium tuberculosis, M. bovis or M. africanum genome. The
CC Mycobacterium tuberculosis, M. bovis or M. africanum genome. The
CC diagnostic reagents comprised within the kit are selected so that they
CC are able to differentiate between M. bovis, M. tuberculosis or M.
CC africanum-infected mammals and mammals vaccinated against M. bovis, M.
CC tuberculosis or M. africanum. The polypeptide or its variant or fragment
CC is useful as a medicament provided that the polypeptide is not a MPT-64
CC polypeptide or a polypeptide encoded by the Rv1984c region of the M.
CC bovis, M. tuberculosis or M. africanum genomes. The nucleic acid which
CC encodes the polypeptide or the diagnostic reagents can be used as a
CC vaccine, especially against Mycobacterium infections. This sequence
CC corresponds to a protein used to derive the peptide epitopes used in the
CC method of the invention.
XX
XX Sequence 100 AA;

Query Match 26.5%; Score 502; DB 7; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.1e-29;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 ATAAQAAYTQAMATTSLPEIAANHITQAVLTATNFFGINTIPALTEMDFYFIRMNQAA 151
 DB 1 ATAAQAAYTQAMATTSLPEIAANHITQAVLTATNFFGINTIPALTEMDFYFIRMNQAA 60
 QY 152 LAMEVYQAEAVNTLFEKLEPMASILDPGASQSTTNPIFG 191
 DB 61 LAMEVYQAEAVNTLFEKLEPMASILDPGASQSTTNPIFG 100

RESULT 13
 ABU34364
 ID ABU34364 standard; protein; 445 AA.
 XX ABU34364;
 XX
 XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #19891.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Mycobacterium bovis.
 OS
 XX WO200277183-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362599P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-0299926/02.
 DR N-PSDB; ACA38234.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 62288; 1766pp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model.
Run on: January 5, 2006, 20:12:50 ; Search time 164 Seconds
(without alignment)
1596.043 Million cell updates/sec

Title: US-09-872-505A-6
Perfect score: 1892
Sequence: 1 MITMLWHAMPPELNTARLWA.....PLAQREREDDDWDEDDW 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1887	99.7	371	Q7D4P4_MYCTU	Q7d4p4 mycobacteri
2	1878	99.3	368	Q7TVG3_MYCBO	Q7tv93 mycobacteri
3	1878	99.3	368	Q7TVG3_MYCBO	Q7tv93 mycobacteri
4	527	27.9	302	O33085_MYCLE	Q33085 mycobacteri
5	445	23.5	435	Q7TXF6_MYCBO	Q7txf6 mycobacteri
6	440	23.3	435	1 PPE47_MYCTU	O53268 mycobacteri
7	336.5	23.1	437	Q7TXF8_MYCBO	Q7txf8 mycobacteri
8	32.5	22.9	518	Q73RJ3_MYCPA	Q73rj3 mycobacteri
9	30.5	22.8	434	1 PPE46_MYCTU	F31500 mycobacteri
10	430.5	22.8	513	Q7U2C9_MYCBO	Q7u2c9 mycobacteri
11	430.5	22.8	513	Q7DA35_MYCTU	Q7da35 mycobacteri
12	426.5	22.5	531	Q73U78_MYCPA	Q73u78 mycobacteri
13	423.5	22.4	504	Q73T11_MYCPA	Q73t11 mycobacteri
14	419	22.1	488	Q741E7_MYCPA	Q741e7 mycobacteri
15	408.5	21.6	550	Q73X20_MYCBO	Q73y20 mycobacteri
16	407.5	21.5	493	Q73D56_MYCPA	Q73td6 mycobacteri
17	404.5	21.4	518	Q7D9S7_MYCTU	Q7d9s7 mycobacteri
18	404.5	21.4	518	Q6MX41_MYCTU	O6mx41 mycobacteri
19	404.5	21.4	518	Q7U1Y9_MYCBO	Q7uly9 mycobacteri
20	400.5	21.2	696	Q741T4_MYCPA	Q741t4 mycobacteri
21	399.5	21.1	538	Q73TF3_MYCPA	Q73tf3 mycobacteri
22	398	21.0	585	Q73WR5_MYCPA	Q73wr5 mycobacteri
23	394.5	20.9	301	Q745B0_MYCPA	Q745b0 mycobacteri
24	383.5	20.3	393	Q79FK6_MYCTU	Q79fk6 mycobacteri
25	383.5	20.3	411	Q7D7Y8_MYCTU	Q7d7y8 mycobacteri
26	382.5	20.2	393	Q7T2J3_MYCBO	Q7t2j3 mycobacteri
27	382.5	20.2	399	Q7TZH8_MYCBO	Q7tzh8 mycobacteri
28	382.5	20.2	403	Q7D7X5_MYCTU	Q7d7x5 mycobacteri
29	382	20.2	536	Q79FY8_MYCTU	Q79fy8 mycobacteri
30	380	20.1	421	Q73ZU6_MYCPA	Q73zu6 mycobacteri
31	380	20.1	536	Q7U2D5_MYCBO	Q7u2d5 mycobacteri

RESULT 1																	
Q7D4P4	MYCTU	Q7D4P4_MYCTU PRELIMINARY; PRT; 371 AA.															
AC	Q7D4P4;																
DT	05-JUL-2004 (TREMBLrel. 27, Created)																
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)																
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)																
DE	PPE family protein.																
GN	OrderedLocusNames=MT3987;																
OS	Mycobacterium tuberculosis.																
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;																
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;																
OC	Mycobacterium tuberculosis complex.																
OX	NCBI_TaxID=1773;																
[1]	NUCLEOTIDE SEQUENCE.																
RP	STRAIN=CDC 1551 / Oshkosh;																
RC	MEDLINE=22206494; PubMed=12218036;																
RX	DOI=10.1128/JB.184.19.5479-5490.2002;																
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,																
RA	Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,																
RA	Hickey E.K., Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M.D.,																
RA	Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,																
RA	Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,																
RA	Fraser C.M.;																
RT	"Whole-genome comparison of Mycobacterium tuberculosis clinical and																
RT	laboratory strains.";																
RL	J. Bacteriol. 184:5479-5490 (2002).																
EMBL	AB000516; AAK48355.1; -; Genomic_DNA.																
DR	TIGR; MT3987; -																
DR	InterPro; IPR000030; Microbac_PPE.																
DR	Pfam; PF00823; PPE; 1.																
SQ	SEQUENCE 371 AA; 37686 MW; CC25F5865131C79B CRC64;																
Query Match 99.7%; Score 1887; DB 2; Length 371;																	
Best Local Similarity 99.7%; Pred. No. 4.3e-98;																	
Matches 370; Conservative 0; Mismatches 1; Indels 0; Gaps 0;																	
QY	1	MITMLWHAMPPELNTARLWAGAGPAPMLAAAGWQTLSAALDAQAVELTARLNSLGEAWT	60														
Db	1	MITMLWHAMPPELNTARLWAGAGPAPMLAAAGWQTLSAALDAQAVELTARLNSLGEAWT	60														
QY	61	GGGSDKALAAATPMVWVLOTASTQAKTRAMQATAQAAAYTQAMATTPLSPLEIAANHITQA	120														
Db	61	GGGSDKALAAATPMVWVLOTASTQAKTRAMQATAQAAAYTQAMATTPLSPLEIAANHITQA	120														
QY	121	VLFTATNFGINTIPIALTENDYFIRMNQAALAMEVYQAEATVNTLFEKLEPMASILDPG	180														
Db	121	VLFTATNFGINTIPIALTENDYFIRMNQAALAMEVYQAEATVNTLFEKLEPMASILDPG	180														
QY	181	ASQSTTNPIFGMSPGSSSTPVGLPPNATQTLQGLGMSGPMQOLTQPLQOQVLSLFSQVG	240														
Db	181	ASQSTTNPIFGMSPGSSSTPVGLPPNATQTLQGLGMSGPMQOLTQPLQOQVLSLFSQVG	240														

Q8vkn6 mycobacteri
Q79fj7 mycobacteri
Q10892 mycobacteri
Q7u2v1 mycobacteri
Q7tzj2 mycobacteri
Q7tzj4 mycobacteri
Q7d7x9 mycobacteri
Q79fk5 mycobacteri
Q7d7y7 mycobacteri
Q79fh3 mycobacteri
Q7d631 mycobacteri
Q79f16 mycobacteri
Q7tz76 mycobacteri
Q7tzr7 mycobacteri

QY 241 GTGGGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPPLMS 300
DB 241 GTGGGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPPLMS 300
QY 301 QLIEKPVAPSVMPAAAGSSATGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERRED 360
DB 301 QLIEKPVAPSVMPAAAGSSATGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERRED 360
QY 361 DEDDDWDEDDW 371
DB 361 DEDDDWDEDDW 371

RESULT 2
QYTVG3 MYCBO
ID Q7TVG3 MYCBO PRELIMINARY; PRT; 368 AA.
AC Q7TVG3
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DE PPE FAMILY PROTEIN.
GN Name=PPE68; OrderedLocusNames=MD3903;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AP2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Manseour H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248347; CAD96089.1; -; Genomic_DNA.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 368 AA; 37330 MW; D78F44095F658CA2 CRC64;

Query Match 99.3%; Score 1878; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-97;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQAKRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQAKRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGQGENSGPMQQLTOPLOQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGQGENSGPMQQLTOPLOQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPPLMSOLI 303
DB 241 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPPLMSOLI 300
QY 304 EKPVPAPSVMPAAAGSSATGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 363
DB 301 EKPVPAPSVMPAAAGSSATGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 360
QY 364 DEDDDWDEDDW 371
DB 361 DEDDDWDEDDW 368

RESULT 4

DB 361 DEDDDWDEDDW 368

RESULT 3
QY9F92 MYCTU
ID Q79F92 MYCTU PRELIMINARY; PRT; 368 AA.
AC Q79F92
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE PPE FAMILY PROTEIN.
GN Name=PPE68; OrderedLocusNames=Rv3873;
OS Mycobacterium tuberculosis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekaita F., Badcock K., Baeham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy J.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; BX842584; CAES5647.1; -; Genomic_DNA.
DR TuberculList; RV3873; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 368 AA; 37330 MW; D78F44095F658CA2 CRC64;

Query Match 99.3%; Score 1878; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-97;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQAKRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQAKRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGQGENSGPMQQLTOPLOQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSPGSGSTPVQQLPPAATQTLGQGENSGPMQQLTOPLOQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPPLMSOLI 303
DB 241 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPPLMSOLI 300
QY 304 EKPVPAPSVMPAAAGSSATGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 363
DB 301 EKPVPAPSVMPAAAGSSATGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 360
QY 364 DEDDDWDEDDW 371
DB 361 DEDDDWDEDDW 368

```

O33085_MYCLE
ID O33085_MYCLE PRELIMINARY; PRT; 302 AA.
AC O33085;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Hypothetical protein MLCB628.14c (PPE-family protein).
GN Name=MLCB628.14c; OrderedLocusNames=ML0051;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93188700; PubMed=8446027;
RA Eigmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT "Use of an ordered cosmid library to deduce the genomic organization
of Mycobacterium leprae."
RL Mol. Microbiol. 7:197-206(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Eigmeier K., Garnier T., De Rossi E., Feihi H., Cole S.T.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churruarin C.M., Harris D.E.,
RA Mungall K.L., Baeham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean A., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
DR EMBL; Y14967; CAA75201.1; -; Genomic DNA.
DR EMBL; AL583917; CAC29559.1; -; Genomic DNA.
DR PIR; T10033; T10033.
DR Leproma; ML0051; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 302 AA; 32135 MW; EDAC4CC0AF3BF3D0 CRC64;

Query Match 27.9%; Score 527; DB 2; Length 302;
Best Local Similarity 45.1%; Pred. No. 6.1e-22;
Matches 134; Conservative 31; Mismatches 78; Indels 54; Gaps 9;

QY 68 LAAATPMVVLQASTQAKTRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLATNF 127
DB 39 IAAAKPMITWLQSA-----AQTTTQAEHQAQMASTPGMAVITENHITQAILATNF 91
QY 128 FGINTPIALTENDYFIRMNQALAMEVYQATAVNTLFEKLEPMASILDPCA----- 181
DB 92 FGINMAPIAFTAGDFICRTOTALAMNSYQAEITLNTAFQKLEPMASILDPCA----- 151
QY 182 --SQ--STTNPIFGMSPGSGSTVQGLPPAATOTLGLGMSGPMQOLTOPLQOVTSLFS 237
DB 152 LTSQVNQFTQMISGSAALPSTVLQ-----QTVGVAE-----LAPMQOVSLFT 198
QY 238 QVGGTGGGNPAD-----EERAAQMLGTSPLSNHPLAGSGSPSAGALLRAESLPAGGSL 293
DB 199 SIDSTGYVTSAGQDIESAHRIGLFGASTLSSHPVLVIGITGTTDTRLLCAESLPASGSL 258
QY 294 TRTPMS--OLIEKPVAPS-----VMPAAAGSSATGGAAPVAGAMGQAQSGGST 343
DB 259 AWTPLMTQFQIDKSTAPEPRQVMPLP-----PWAAGSPGHNAQDGGTT 302

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 302 AA; 32135 MW; EDAC4CC0AF3BF3D0 CRC64;

RESULT 5

```

```

Q7TXF6_MYCBO
ID Q7TXF6_MYCBO PRELIMINARY; PRT; 435 AA.
AC Q7TXF6;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE PPE FAMILY PROTEIN.
GN Name=PPE47; OrderedLocusNames=Mb3047c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248344; CAD96734.1; -; Genomic DNA.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 435 AA; 42878 MW; 62A5615EB0D0DBED CRC64;

Query Match 23.5%; Score 445; DB 2; Length 435;
Best Local Similarity 34.0%; Pred. No. 3.6e-17;
Matches 136; Conservative 50; Mismatches 148; Indels 66; Gaps 14;

QY 1 MTMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGE-AW 59
DB 1 MTPAVMLASPPPEVSHALLSAGPGPSLQAAAAGNSALSABYAQAQELSAVVAAGVGV 60
QY 60 TGGGSDKALAAATPMVVLQASTQAKTRAMQATAQAAYTQAMATTPSLPEIAANHITQ 119
DB 61 QPSAELFVAAYVYVAVMLVQASADSAAAAAGEHEAAAGVVCALAEMPTLPALAANHLTH 120
QY 120 AVLATNPFINTPIALTENDYFIRMNQALAMEVYQAT-AVNTLFEKLEPMASILD 178
DB 121 AVLATNPFINTPIALTENDYFIRMNQALAMEVYQAT-AVNTLFEKLEPMASILD 179
QY 179 PGASQSTTNPIFGMSPGSGSTVQGLP---PAATOTLGLGEM-----SGPMQOLT 226
DB 180 PGANEASNAVAATITP---FPFGLAKFLEMAQAFTVEGELIMKSAEAWAVGFVELIT 236
QY 227 ----QPLQQVTSLS---PSQVG---GTGGGNPADEEAAQM-----GLLGTSP 262
DB 237 GLVNFPEMLVLTGMIDMFFATVGFALGVFLVPLLEFAVVLLELAILLSIGWIIISNIFGAIP 296
QY 263 LSNHPLAGSGSPSAGAGLLRAESLPAGGSLTRTPMSQLIEKPVAPSVMPPAAAGSSAT 322
DB 297 VLAGPLLGA-----LAAAVVPGVAG-----VTGLAGLAAPVAVGAAGAPAAV 340
QY 323 GGNAPVAGAMGQAQSGSGSTRPLVAPAP-----LAQER 357
DB 341 GSVAPVGGVSPQARLVSAVEP---APASTSVSVLSADR 377

RESULT 6
PPE47_MYCTU
ID PPE47_MYCTU STANDARD; PRT; 435 AA.
AC O53268; O53269;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical PPE-family protein PPE47/PPE48.
GN Name=ppe47/ppe48; OrderedLocusNames=RV3022c; MT3106;
OX ORFNames=MTV012.35c/MTV012.36c;
OS Mycobacterium tuberculosis.

```



```
Db 303 PILGA-----LAAAVVPGVAG-----LAGVAGLAAPVAVCAAGAPALVGSVA 346
QY 327 PVGAGMGGAQSGGSTRGLVAPAP-----LAQER 357
Db 347 PVSGGVSPQARLVSAVEP---APASTSVSLASDR 379

RESULT 8
Q73TJ3 MYCPA
ID Q73TJ3 MYCPA PRELIMINARY; PRT; 518 AA.
AC Q73TJ3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAF3725;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017240; AAS06275.1; -; Genomic_DNA.
DR InterPro; IPR00030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 518 AA; 52170 MW; BAC863B26F9CCF62 CRC64;

Query Match 22.8%; Score 432.5; DB 2; Length 518;
Best Local Similarity 32.6%; Pred. No. 2.2e-16;
Matches 143; Conservative 41; Mismatches 155; Indels 99; Gaps 15;

QY 1 MTTMLHMPPELNTARLMAGAPAPMLAAAGCQTLASALDAQAVELTARLNSL-GEAW 59
Db 1 MTAPIWMAFPEVHSAALLSGPGGLASAGAWNSLALEYTSAAELTAVGCVQAGAW 60
QY 60 TGGGSKALAAATPMVWVLTQASTQAKRAMQATAQAAAYTQAMATTPSLPEIAANHITQ 119
Db 61 QGSAESYVAHAPYLAWLTQASDSAGAAARHEAAAAAYTALAAMPTLGEIAANHATH 120
QY 120 AVLTATNFFGINTPIALTEMDFEIRMNQALAMEVYQABTAVNTLFEKLEPMASILD 179
Db 121 AALQATNFFGINTPIAIVNEADY-ARMWQAAATTTTYQ---AVSTAATAVATPQ-TMPAP 175
QY 180 GASQSTNPI-FGMPSPGSS-----TPVGQLPPAA-TQ-----TL 212
Db 176 QIAKSTAATIPFORPVNPTNFSQVWADLWYDIPYSIETFGFVDPANWTQLFQFVNVT 235
QY 213 GOLGEMSGPMQO-----LTQPL-----Q 230
Db 236 ANLAGTPAKAQLFSPNSVLFSPNLTLLWLDFTAGRIFDLVTILKFLLEQLLYVVLG 295
QY 231 QVTSLSFQVGGTGGGNPADEAAQMGILGTSPISNHPLAGG-----SGPSAGALLRAE 284
Db 296 AVTSLGAAAGAAG-----LVGLAGLA--SPGLPTGAEMVPTVAPPPTGATPAPT 344
QY 285 SLPGAGSLTRTPLMSQLIEKVPAPVMPA-AAAGSSATGGAAP-----VGAGMGQGA 337
Db 345 FLIGGPASAPATIPVSSAAAAAPTAPAPAPAAVAAAGSAAPPFGGPGFPYLVGGRVRS 404
QY 338 QSGGSTPGLVAPAPLAQ 355
Db 405 SASAQSRPKSAASAAAAA 422

RESULT 9
PPE46 MYCTU
ID PPE46 MYCTU STANDARD; PRT; 434 AA.
AC P31500; O53265;
```

```
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical ppe-family protein PPE46.
GN Name=ppe46; OrderedLocusNames=Rv3018c, MT3098/MT3101;
GN ORFNames=MIV012.32c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekla F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornaby T., Jagsis K., Krogh J., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischnmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "whole-genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 160-374.
RC STRAIN=Isolate 50410;
RA Patki A.H., Dale J.W.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the mycobacterial ppe family.
CC -!- CAUTION: In strain Oshkosh the gene for this protein is
CC interrupted in position 307 by an IS6110 element.
CC -!- CAUTION: Was originally (Ref.3) thought to be a dihydrofolate
CC reductase.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC in positions 294, 337 and 355.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX842581; CAE55538.1; -; Genomic DNA.
CC EMBL; AE000516; AAK47427.1; ALT SEQ; Genomic DNA.
CC EMBL; AE000516; AAK47430.1; ALT SEQ; Genomic DNA.
CC EMBL; X59271; CA41961.1; ALT_FRAME; Genomic DNA.
CC PIR; E70857; E70857.
CC TIGR; MT3098; -.
CC TIGR; MT3101; -.
CC TubercuList; Rv3018c; -.
CC InterPro; IPR00030; Microbac_PPE.
CC Pfam; PF00823; PPE; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 434 AA; 43029 MW; 41D673C4BD389DD6 CRC64;

Query Match 22.8%; Score 430.5; DB 1; Length 434;
```


Query Match 22.8%; Score 430.5; DB 2; Length 513;
 Best Local Similarity 33.3%; Pred. No. 2.8e-16;
 Matches 134; Conservative 44; Mismatches 155; Indels 69; Gaps 13;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVALTARLNSL-GEAW 59
 DB 1 MAAPITWASPPPEVHSALLSNGPGSLVAATAWQSLSAEYASTAELSGLLGAVPGWAW 60

QY 60 TGGGSDKALAAATPMVWMLQTAQKTRAMQTAQAATAATTAATPSLPEIAANHITQ 119
 DB 61 QGSAEYVAHLPYVAVMLTQASADAAGAAQHEAAAAAYTTALAAMPTLAELAAAHVH 120

QY 120 AVLTATNPFINTIPALTMDYFIRMWQALAMEVYQAET-AVNTLFEKLEPMASILD 178
 DB 121 TVLVATNPFINTIPALTMDYFIRMWQALAMEVYQAET-AVNTLFEKLEPMASILD 178

QY 179 PGASQSTTNPIFGMPSPGSSSTP-----VGLPPAATQTLQGLGEMSGPMQO----- 224
 DB 180 FGGGAAT-----VGVNFWQWLLALLQQLWNNYTGFGYGMWLLQIWLQDPIGNSI 231

QY 225 -----LTOPLOQVTS-----LFSQVGGTGGGNPADEEAAQMGLLGTSPLSNH 266
 DB 232 KIILAFITNPIQALITYGELLFALGYOIFPNLVGM-----PTWGMILSSPFL-L 279

QY 267 PLAGGSGPSAGAGL-----LRAESLPGAGGSLTRTPMLSQLIEKPVAPSVMPAAAGSSATG 323
 DB 280 PAGLGLGLAAIAFLPVLPAVIPPAS-----TPLAAAV-----AAGSVMPAV---SMAVT 328

QY 324 GAAPVAGAGAGGAGGSGSTRPGLVAPAPLAQEREDDDDD 365
 DB 329 GAGTAGAATPAGAAPSAAGAPAPATASAYAVGGSGDW 370

RESULT 12
 Q73U78 MYCPA
 ID Q73U78 MYCPA PRELIMINARY; PRT; 531 AA.

AC Q73U78;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MAP3490;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=k10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017239; AAS06040.1; -; Genomic_DNA.
 DR InterPro; IPR000030; Microbac_PPE.
 DR Pfam; PF00823; PPE; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 531 AA; 53742 MW; B392FE119582DC8C CRC64;

Query Match 22.5%; Score 426.5; DB 2; Length 531;
 Best Local Similarity 34.9%; Pred. No. 4.9e-16;
 Matches 137; Conservative 35; Mismatches 161; Indels 59; Gaps 15;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVALTARLNSL-GEAW 59
 DB 7 MTAPIWASPPPEVHSALLSNGPGSLVAATAWQSLSAEYASTAELSGLLGAVPGWAW 66

QY 60 TGGGSDKALAAATPMVWMLQTAQKTRAMQTAQAATAATTAATPSLPEIAANHITQ 119
 DB 67 QGSAEYVAHLPYVAVMLTQASADAAGAAQHEAAAAAYTTALAAMPTLAELAAAHVH 126

QY 120 AVLTATNPFINTIPALTMDYFIRMWQALAMEVYQAET-AVNTLFEKLEPMASILD 178

DB 127 GALVATNPFINTIPALTMDYFIRMWQALAMEVYQAET-AVNTLFEKLEPMASILD 185

QY 179 PGASQSTTNPIFGMP-----SPGSSSTPVGQLPPAATQTLG-----QLGEMSG-PMOQLTQPL 229

DB 186 STAAHDDHDEHGDHDDHDFDPLNQFVAQLRLFGIDWDVPVETGLNGLPYEAYTSPA 245

QY 230 Q-----QVTSLSFSQVGGTGGGNPA-----DEEAAQW-GLLGTSP-LS 264

DB 246 DPLMWVVRALBELFDFQFGALLQENPAAAFQITELVLLDMPTHLAQLASWLPQOQL 305

QY 265 NHPLAGGSGPSAGAGLRAESLPGAGGSLTRTPMLSQLIEKPVAPSVMPAAAA---GSSAT 322

DB 306 LVPALVAAPFG-----ALAGPAGVAGQPPLEPA-----PVAEPATPSAAAPTGLPAT 352

QY 323 GGAAPVAGAGAGGAGGSGSTRPGLVAPAPLA 354

DB 353 AGATPIAASA-----AASGPAPAP---TPAPTA 377

RESULT 13
 Q73T11 MYCPA
 ID Q73T11 MYCPA PRELIMINARY; PRT; 504 AA.

AC Q73T11;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MAP3737;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=k10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017240; AAS06287.1; -; Genomic_DNA.
 DR InterPro; IPR000030; Microbac_PPE.
 DR Pfam; PF00823; PPE; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 504 AA; 49761 MW; 4DD3E220EB495932 CRC64;

Query Match 22.4%; Score 423.5; DB 2; Length 504;
 Best Local Similarity 32.9%; Pred. No. 6.8e-16;
 Matches 138; Conservative 39; Mismatches 140; Indels 103; Gaps 13;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVALTARLNSL-GEAW 59
 DB 1 MTAPVWASPPPEVHSALLSNGPGSLVAATAWQSLSAEYASTAELSGLLGAVPGWAW 60

QY 60 TGGGSDKALAAATPMVWMLQTAQKTRAMQTAQAATAATTAATPSLPEIAANHITQ 119
 DB 61 EGPTTEAYVAHLPYVAVMLTQASADAAGAAQHEAAAAAYTTALAAMPTLAELATNATH 120

QY 120 AVLTATNPFINTIPALTMDYFIRMWQALAMEVYQAET-AVNTLFEKLEPMASILD 177
 DB 121 AALVATNPFINTIPALTMDYFIRMWQALAMEVYQAET-AVNTLFEKLEPMASILD 176

QY 178 -----DPGAS-----QSTTNPIFGMPSPGSSSTPVGQLPPAAT 209

DB 177 IVKSESSAGDTGDSGGGGFGDLLSAWENFVWMLMDQLFGVDS-----PDLAS 227

QY 210 QTLGQGE-----MSGPMOQLTOPLOQVTSLSFSQVGGTGGGNPADEEAAQMGLLGTSPLS 264

DB 228 AVAAFLANPSALFALLFALAYEIAFDLFFSPV-----ALLAAPFLP 271

QY 265 NHPLAGGSGPSAGAGLRA-----BSLPGAGGSLTRTPMLSQLIEK 305

DB 272 FVGLAGLAGLAGLAGAGAPAPVAVPAPTSPPNQALPVAGLTPSASTPAS----- 325

QY 306 FVAPSVMPAAAAAGSSATGGA-APVGAGM-----GGAQSGGST--RPLVAPAPLA 354

```
Db      326 --APASAPATSATAPASAPAPAGSPVFGYLVGGGSGTTLTGSRSNATAPAGLA 383
|||:||:::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
RESULT 14
Q741E7_MYCPA
ID   Q741E7_MYCPA PRELIMINARY;          PRT;    488 AA.
Q741E7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=MAP1144c;
OS   Mycobacterium paratuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium avium complex (MAC).
NCBI_TaxID=1770;
[1]
RN   NUCLEOTIDE SEQUENCE.
RP   STRAIN=X10;
RA   Li L., Barnattant J., Zhang Q., Anonsin A., Alt D., Kapur V.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AE017230; AAC03461.1; -; Genomic_DNA.
DR   InterPro; IPR000030; Microbac_PPF.
DR   Pfam; PF00823; PPE; 1.
KW   Complete proteome.
SQ   SEQUENCE 488 AA; 48598 MW; C2FD0FB977D3334B CRC64;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2006, 16:44:30 ; Search time 40 Seconds
(without alignments)
892.410 Million cell updates/sec

Title: US-09-872-505A-6
Perfect score: 1892
Sequence: 1 MITMLWHAMPPELNTARLMA.....PLAQRERDEDDWDEDDW 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	99.3	368	2 G70802	probable PPE prote
2	527	27.9	302	2 T10033	hypothetical prote
3	430.5	22.8	434	2 E70857	probable PPE prote
4	430.5	22.8	513	2 D70836	probable PPE prote
5	404.5	21.4	518	2 F70831	probable PPE prote
6	383.5	20.3	333	2 G70929	probable PPE prote
7	382	20.2	536	2 F70835	probable PPE prote
8	379.5	20.1	403	2 H70931	probable PPE prote
9	376.5	19.9	463	2 H70750	probable PPE prote
10	376	19.9	423	2 B70931	probable PPE prote
11	374	19.8	350	2 H70929	probable PPE prote
12	373	19.7	473	2 F70813	probable PPE prote
13	371	19.6	391	2 D70922	probable PPE prote
14	369.5	19.5	385	2 H70503	probable PPE prote
15	365.5	19.3	391	2 A70663	probable PPE prote
16	365	19.3	556	2 D70940	probable PPE prote
17	362	19.1	408	2 G70925	probable PPE prote
18	360.5	19.1	413	2 F70560	probable PPE prote
19	356	18.8	572	2 T45392	hypothetical prote
20	352	18.6	391	2 B70625	probable PPE prote
21	351.5	18.6	365	2 E70929	probable PPE prote
22	348	18.4	463	2 C70931	probable PPE prote
23	344.5	18.2	394	2 A70504	probable PPE prote
24	343.5	18.2	380	2 A70446	probable PPE prote
25	331	17.5	468	2 B87158	PPE-family protein
26	326	17.2	423	2 C70582	probable PPE prote
27	325.5	17.2	438	2 S77655	hypothetical prote
28	323.5	17.1	468	2 B70932	probable PPE prote
29	322.5	17.0	406	2 E70675	probable PPE prote

ALIGNMENTS

RESULT 1

G70802
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70802
R:Coles, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, R.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70802
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-368 <COL>
A:Cross-references: UNIPROT:O69738; UNIPARC:UPI0000165388; GB:AL022120; GB:AL123456; N1:
A:Experimental source: strain H37RV
C:Genetics: PPE
A:Gene: PPE

Query Match	99.3%	Score 1878;	DB 2;	Length 368;
Best Local Similarity	100.0%	Pred. No. 1.9e-104;		
Matches	368;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
QY	4	MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG	63	
Db	1	MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG	60	
QY	64	SKALAAATPMVWLOTASTOAKTRAMQATQAAYTOAWATTPSLPEIAANHITQAVLT	123	
Db	61	SKALAAATPMVWLOTASTOAKTRAMQATQAAYTOAWATTPSLPEIAANHITQAVLT	120	
QY	124	ATNFFGINIPIALTENDYFIRWNOALAMEYVQAEVAVNTLFEKLEPMASILDPCASQ	183	
Db	121	ATNFFGINIPIALTENDYFIRWNOALAMEYVQAEVAVNTLFEKLEPMASILDPCASQ	180	
QY	184	STTNPIFGMPSPGSGSTFVQQLPPAATQTTLGQMGSGPMQQLTQPLQQVTSLSFQVGGTG	243	
Db	181	STTNPIFGMPSPGSGSTFVQQLPPAATQTTLGQMGSGPMQQLTQPLQQVTSLSFQVGGTG	240	
QY	244	GNPADERAAQMGLLGTSTPLSNHPLAGGSPSAGALLRAESLPGAGGSLTRTPPLMSQLI	303	
Db	241	GNPADERAAQMGLLGTSTPLSNHPLAGGSPSAGALLRAESLPGAGGSLTRTPPLMSQLI	300	
QY	304	EKVPASVMPAAAGSSATGGAPVCGAGAGQAGGSGTRPGLVAPAPLAQERREDDDED	363	
Db	301	EKVPASVMPAAAGSSATGGAPVCGAGAGQAGGSGTRPGLVAPAPLAQERREDDDED	360	
QY	364	DWDEDDW 371		


```

Db      276 FDFPTHMLQLATWLAEENPLLVAALTPAISGLGAVSGLAGLITGLVPQPPVVVPAAPDAVV 338
        Qy    309 PSWNPAAAASSSATGGNAFPVGACAMGCQAQGSGSTRFLGVAP 350
            |:|:| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
        Db     336 PTVLPFLAGTATTPTTAASAPAAGA-APGPPPAGTAGTATSASVP 376


RESULT 6
G70929
probable PPE protein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70929
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;
Hartung, M.P.; Kaul; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Hoiczyk, A.;
Karpman, A.; Kumarasinghe, T.M.; Leung, K.Y.; MacIntyre, S.; Madico, G.; Nagarajan, N.; Narasimhan, B.;
Nayak, U.V.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.;
Struelens, L.S.; Taylor, W.H.; Whitehead, S.; Whittam, T.S.; Young, L.D. et al., Nature 393, 537-544, 1998
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence
A:Reference number: A70500; UID:98295987; PMID:9634230
A:Accession: G70929
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-393 <COL>
A:Cross-references: UNIPROT.O53939; UNIPARC:UIO0001652AA; GB.AL022021; GB.AL12
A:Experimental source: Strain H37Rv
A:Genetics:
A:Gene: ppe
```

```

184 AAVATAAGTAQSTLTTEMITGLPN-----ALQSLTSPLLQSSNGSPLSWLDWQLL 233
QY 228 -----PLQVTSLSFQVG-----TGGCNPADEEAQAQGLLCTSPLSNHP 267
D 231 FGTNPFTSISALLTDLQFYASFFYNTTEGLPYFSIGMNNFTQSAKTGLGLGSAAPA 290
QY 268 LAG-----GSGPSAGAGLLRAESLPGAGGSLTRTFLMSQLIEKVP----- 309
D 291 AAGDAAKGLPLGGMLGGGPVA-AGLGNAAV-----GKLSVDPVWSGPLGSVTPGCAAPL 345
QY 310 -----SVMPAAAGSSATGGAAPVCGAGMCGAQQSGGSTRPGLVAPAPLA 354
D 346 PVSTVSAAPAPAPG-SLLGGLPLAGAGGAGAPRYG--PRPTVMARPPPA 392

RESULT 7
F70835
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70835
R: Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;
  Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.;
  Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.;
  Squares, S.; Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sullston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.

```


Qy	9	MPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARINSI-GEAWTCGGSDKA	67
		: :	
Db	10	LPPEINARMYSGAGSAPMLAAASAMHGLSABLRASALSYSVLSTLTGEEHWGPASASM	69
		: :	
Qy	68	LAAATPMVNLQTASTOAKTRMCAQAARAAYTOAMATPSLPEIRANHITOAQLVTATNF	127
		: :	
Db	70	TAAAPYVMWSVTVAVRREQAGAEAAAAYEAFATVPFVIEANRAQLMALIAITNV	129
		: :	
Qy	128	FGINTIPIALTEMDYFIIRMNQALAMEVYQ-AETAVENTIFKELEPW-----	173

Db	130	LGQWAPAIATEAQY-AEMWSQDAMMYGAGASAAATQLTPPEPQTINASGLAAQSA	188
Qy	174	-----ASILDPGASQSTNPP-IFQWPS--PG-SSTPVGQLPPAATQTTLGQLGEMSGPMQQL	225
Db	189	AIAHATGASGAAQQTTLISQLIAATPVSVLQGLSSSTATFASGPGSLGLIVGSGSSWLDKL	248
Qy	226	TQPLQQTVSFLFSOVGGTGGGNPADERAAOMGLLG-----TS	261
Db	249	WALLDPNSNTFWNTIASSGLFLPENTTAPFLLGLGGVAAADAAGDVLGEATSGGLGGALVA	308
Qy	262	PLSNHPIAGSGPGSAGAGLLARLSLPGAGSLTRTPLMSQLIEKPVAPSVWMPAAAGSSA	321
Db	309	PLGS-----ACGLGGTVAAGLGNAAATV-----GTLSPVPSWT-----AAAPLASPLGSA	352
Qy	322	TGG---AAP---VCAGAMG-----QGAQSGGST-----RPEGLVAPAPLA	354
Db	353	LGGTPTWVAPPVPAVNAAGMPGMPFGTMGGQGFGRAVPQYGFNPVARPAA	402
RESULT 9			
H70750			
Probable PPE protein - Mycobacterium tuberculosis (strain H37Rv)			
C/Species: Mycobacterium tuberculosis			
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004			
C/Accession: H70750			
C/Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;			
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Ho-			
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.			
Nature 393, 537-544, 1998			
A/Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.			
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complet			
A/Reference number: A70500; MUID:9825987; PMID:9634230			
A/Accession: H70750			
A/Status: preliminary; nucleic acid sequence not shown; translation not shown			
A/Molecule type: DNA			

[illegible]

181	TRDVLSSIGQLIRDIILDFIANPKYKFLFFFEQGF-SPAVTVVLVALVALQLYDFLWTPY	239
Db		
203	-----QLPFAATQTLGLGEMSGPMQQLTQLPQQVTSLFSQVGGTGGGNPADEEAAQMG	256
QY		

Db 240 YASYGLLLPFPPTLSALTALSLHLNLPAGLLPAAALG-----PGDQGANLA 293
 QY 257 LGTSPSLNHPLAGSGPAGAGLRAESLPAG--GSUTRPLMSQLIEKPVASVMPA 314
 Db 294 VAVTATAAHP--GGSPTSPNP-----APAAPSSNSVGSASAPGISYAVFGLAPGVSSG 347
 QY 315 AAGSSATGGAAPVGAAGMGQAGQSGSTRPGL 347
 Db 348 PRAGTKSPDTAADTLATA-----GAARPG 372

RESULT 10
 B70931
 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: B70931
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: B70931
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-423 <COL>
 A:Cross-references: UNIPROT:O53950; UNIPARC:UPI00000D5E6B; GB:AL022021; GB:AL123456; NID:123456
 A:Experimental source: strain H37RV
 C:Genetics: PPE
 A:Gene: PPE

Query Match 19.9%; Score 376; DB 2; Length 423;
 Best Local Similarity 31.7%; Pred. No. 2.5e-15;
 Matches 128; Conservative 35; Mismatches 153; Indels 88; Gaps 11;
 QY 9 MPPELNTARLMAGAGPAPMLAAAGWQTLTSAALDAQAVELTARLNSLGEAWTGGSDKAL 68
 Db 6 LPPEINSGRMVTFGPGPMLAAATAWDGLAVELHATAAGYASBELSALTGAWSGPSSMSA 65
 QY 69 AATPMVWLQTAQKTRAMQATQAQAAAYTQAMATTPSLPEIAANHITQAVLTATNF 128
 Db 66 SAAAPYVAMMSATVAELAGRAQLAIAAYEAFAATVPPVIAANRAQLMVLITATNF 125
 QY 129 GINTIPALTENDYFIRMWNOALAMEVYQ----- 158
 Db 126 QNTPALIMTEAQY--MEMWAQDAAMVGYAGSSATASRTAFTEPPQTTHHGQLGQSSA 184
 QY 159 -AETA-----VNTLFEKL-----EPMASILDGASQSTT-----NPIFGMP 193
 Db 185 VAQTAATAAGNLQSAFFQLLSAVPRALQGLALPTASQSASATPQWVTDLGNLSLFLGGA 244
 QY 194 SPSSSTPVGQLPAAATQTLQGLGEMSGPMQOLTQPLQVTSLSQVGG---TCGGNPADE 250
 Db 245 VTGPYTPPGVLPSPGVPYL--LGTSQ---VLVTQNGQGVSAALGKIGKPKITGALAPLAE 299
 QY 251 EAAQMGILGTSPLSNHPLAGSGPSAG---AGLLRAESLPAGGSLTRTPLMSQLIEKPV 307
 Db 300 FALHTPILGSEGL-----GGGSVSAIGRAGLVKLSVPGW-----TV 338
 QY 308 APSVMPAAAAAGSSATGGAAPVGAAGMGQAGQSGSTRPGLVAPA 351
 Db 339 AAPEIFSPAAALQATRLAABAPIAATDAGALGGMALSGLAGRA 382

RESULT 11
 H70929
 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: H70929

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70929
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-350 <COL>
 A:Cross-references: UNIPROT:O53940; UNIPARC:UPI00001652AB; GB:AL022021; GB:AL123456; NID:123456
 A:Experimental source: strain H37RV
 C:Genetics: PPE
 A:Gene: PPE

Query Match 19.8%; Score 374; DB 2; Length 350;
 Best Local Similarity 32.9%; Pred. No. 2.6e-15;
 Matches 124; Conservative 42; Mismatches 149; Indels 62; Gaps 13;
 QY 8 AMPELNTARLMAGAGPAPMLAAAGWQTLTSAALDAQAVELTARLNSLGEAWTGGSDK 66
 Db 5 ALPPEINSGRMVTFGPGPMLAAATAWDGAVELGLATGYASVIAELTGAPVWGAASLS 64
 QY 67 ALAAATPMVWLQTAQKTRAMQATQAQAAAYTQAMATTPSLPEIAANHITQAVLTATN 126
 Db 65 MVAATPYVAMLSQAAARAEQAGMQAAAAAAYEAFAVMTVPPVITANRVLMTLIATN 124
 QY 127 FPGINTIPALTENDYFIRMWNOALAMEVYQATVNTLPEKLEPMAS-----I 176
 Db 125 FFGQNSAALAVAEAOY--AEMWAQDAVAMVGYAASASAS---RLIPFAAPPKTTNSAGVV 180
 QY 177 LDPGASQSTNP-----IFGMPSGSSSTPVGQLPAAATQTLGQGENSGPMQQLTQPL 229
 Db 181 AQAVASVSWPNENDMWLVRLGSLITPTTIVRL-----LQGSYLATGMARFLTISA 233
 QY 230 QQVTSLSFGSGVGTGGGNPADEEAAQMLLGTSPLSNHPPLAGSGPSAGAGLLRAESLPGA 289
 Db 234 QQLT--FGPGITAGSGGAWYTPPQAGL-----GAGPAVSASLARAEPV--- 276
 QY 290 GGSLSLTPMLSQLIEKPVAPSVMPAAAAAGSSATGGAAPVGAAGMGQ---AQSGGST--- 343
 Db 277 -GRLSVPP--SWAVAAP-APAEKPEAGTSPMSVITGEASCGQGLLRGILPLARAGRRTGAF 332
 QY 344 -----RPGIVAPAPLA 354
 Db 333 AHRYGFRHSVITRSPSA 349

RESULT 12
 F70513
 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: F70513
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70513
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-473 <COL>
 A:Cross-references: UNIPARC:UPI00000C1594; GB:297559; GB:AL123456; NID:G3261820; PIDN:C
 A:Experimental source: strain H37RV
 C:Genetics: PPE
 A:Gene: PPE

Query Match 19.7%; Score 373; DB 2; Length 473;
 Best Local Similarity 28.9%; Pred. No. 4.2e-15;

C;Accession: A70663
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: A70663
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-391 <COL>
A;Cross-references: UNIPROT:P95247; UNIPARC:UPI00000D5E93; GB:Z83860; GB:AL123456; NID:9
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: PPE

Query Match	19.3%;	Score 365.5;	DB 2;	Length 391;
Best Local Similarity	30.3%;	Pred. No. 9.5e-15;		
Matches 120;	Conservative 38;	Mismatches 165;	Indels 73;	Gaps 12;

QY	9	MPPELNTARLMAGAGPAPMLAAAAGWQTLSAALDAQVELTARLNSLGEA-WTGGSDKA	67
Db	8	LPPEINSARIYAGAGGGLPFMAAAAWEGLAIDLRSASSFDVIAAGLAAGPWSGPAVSAM	67
QY	68	LAATPMVWVLOTASTQAKTRAMQATAQAAATQAMATTPSLPEIAANHITQAVLTATNF	127
Db	68	AGAAAAPYVGLSAAAGQAELSAQATAAATFEEALAAATVHPAAVTANRVLGALVATNI	127
QY	128	FGINTIPIALTMDYFIRMNQAALAMEVYQAEAVNTLFEKLEP-----	172
Db	128	LQONTPAIAATEFDY-VEWMAQDVGMVGYHAGAA--AVAETLTSPSPPLDLAGLASQA	184
QY	173	-----MASILDPGASQSTTNPIFGMPS--PGSSTPVGQLPPAAQTTLGOLGE-----MS	219
Db	185	GQLTGMATSVGAALSPIAEGAVEGVPAVVAAAQSVAAGLPVDAAALQVQAAAAYPASMLI	244
QY	220	GPMMQITQPLQQTSLFSQVGGTGGGNPADEEAAQMGILLGTSPLSNHPLAGSGSPSAGAG	279
Db	245	GPMMQLAQ-----MGTTANTAGLAGAAGAAAGLAADV-----PTFAGDIASGTGLGGAGG	293
QY	280	LLRAESLPGAGGS--LTRTPMSQLTEKVPAPSPVMPAAAAAGSSATGGAA-----	326
Db	294	L-----GAGMSAEIGKARLVGAMSVPTTWEGSVPARMASSAMAGLGMPAEVPAAGGP	346
QY	327	-----PVGAGAMGCGAQS-----GSTRPGLVAPAP	352
Db	347	MGMMPMPMGNGGAGAGAGPAGMGRGGGANPHVVQARP	382

Search completed: January 5, 2006, 21:15:37
Job time : 43 secs

This Page Blank (uspro)